

A Mixed Basis Density Functional Approach for Low Dimensional Systems with B-splines

Chung-Yuan Ren^{a,†}, Chen-Shiung Hsue^b, and Yia-Chung Chang^{c,d}

^a *Department of Physics, National Kaohsiung Normal University, Kaohsiung 824, Taiwan*

^b *Department of Physics, National Tsing Hua University, Hsinchu 300, Taiwan*

^c *Research Center for Applied Sciences,
Academia Sinica, Taipei 115, Taiwan*

^d *Department of Physics, National Cheng-Kung University, Tainan, Taiwan 701*

[†] *E-mail address: cyren@nknknu.edu.tw*

Abstract

A mixed basis approach based on density functional theory is employed for low dimensional systems. The basis functions are taken to be plane waves for the periodic direction multiplied by B-spline polynomials in the non-periodic direction. B-splines have the following advantages: (1) the associated matrix elements are sparse, (2) B-splines possess a superior treatment of derivatives, (3) B-splines are *not* associated with atomic positions when the geometry structure is optimized, making the geometry optimization easy to implement. With this mixed basis set we can directly calculate the total energy of the system instead of using the conventional supercell model with a slab sandwiched between vacuum regions. A generalized Lanczos-Krylov iterative method is implemented for the diagonalization of the Hamiltonian matrix. To demonstrate the present approach, we apply it to study the C(001)-(2×1) surface with the norm-conserving pseudopotential, the n-type δ -doped graphene, and graphene nanoribbon with Vanderbilt's ultra-soft pseudopotentials. All the resulting electronic structures were found to be in good agreement with those obtained by the VSAP code, but with a reduced number of basis.

PACS: 71.15.Mb, 73.20.-r

Keywords: Density functional theory, a mixed basis, B-splines, low-dimensional systems

I. INTRODUCTION

The electronic properties of low-dimensional systems are fundamentally different from those in higher dimensions due to their unusual collective excitations. Nowadays, 1-dimensional (1D) or 2-dimensional (2D) materials can be easily fabricated due to the emerging nanotechnology, which leads to intensive exploration of low-dimensional systems for material innovation.

First-principles methods based on the density functional theory have been used extensively in investigating the electronic structures and properties of solids. In most calculations, three-dimensional (3D) plane waves are used as basis functions, which are very suitable for systems that are periodic in three dimensions. Because of its easy implementation and the fact that the convergence of the calculation can be checked systematically, plane waves are often employed to expand the wavefunction even along the non-periodic direction in low-dimensional systems via the use of supercell. In this approach, the physical low-dimensional system is treated as a fully 3D periodic system by introducing some artificial vacuum space to separate the periodic replica along the direction, in which the system should be considered as nonperiodic (Fig. 1(a)). The drawback of this approach is the requirement of large thickness of the vacuum layer such that the interactions between the adjacent slabs are negligible, and therefore increases the number of the plane wave along that direction. More seriously, for charged systems (e.g. charged defects), the rather long-range tail of the Coulomb potential inevitably requires an extremely large separation of the two slabs and makes the calculation impractical.

Li and Chang have previously introduced a mixed planar basis which is the product of 2D plane waves in the periodic directions and 1D Gaussian function along the growth direction z for the first-principles calculations [1]-[3]. The use of this mixed basis resumes the layer-like local geometry which appears in surfaces. This mixed basis is suitable for describing electronic properties of low-dimensional systems including surfaces, interface, and superlattices. Moreover, because one can calculate the total energy for an isolated slab instead of using a supercell consisting of alternating slab and vacuum regions, the physical quantity of a surface, such as the work function can be immediately obtained without any correction.

However, the Gaussian-type orbital basis falls off too rapidly for large z , i.e., it behaves

like e^{-z^2} rather than e^{-z} for the hydrogenic orbitals. Increasing the number of the orbitals makes no guarantee of numerical stability and good convergence. Furthermore, the Gaussian orbital is an atom-centered basis and will move accordingly in geometry optimization process, making the coding more difficult and harder to maintain. Inspired by previous works [4]-[6], in the present work we replace the Gaussian orbitals by B-splines [7]. B-splines are highly localized and piecewise polynomials within prescribed break points which consist of a sequence numbers called knot sequence. B-splines have several advantages over Gaussian basis functions: (1) the relevant matrices are sparse. (2) B-splines are superior to traditional methods in the treatment of derivatives. (3) B-splines possess good flexibility to represent a rapidly varying wavefunction accurately with the knots being arbitrarily chosen to have an optimized basis. (4) Finally, B-splines are independent of atomic positions, so the geometry optimization can be easily implemented.

We also implement a generalized Lanczos-Krylov iterative method for the diagonalization of the Kohn-Sham Hamiltonian. It offers a stable and promising way to find the low-lying eigenvalues of a real matrix with a very large dimension via an iterative scheme.

To test the application of the present approach, we will choose the C(001)-(2×1) surface, n-type δ -doped graphene, and the graphene nanoribbon as three examples. We perform the band structure calculation for the former case using the norm-conserving pseudopotential (NCPP) [8]. For the latter two cases, we adopt Vanderbilt's ultra-soft pseudopotentials (USPP) [9] with modest energy cutoff. It is found that the band structures are all in good agreement with those by the popular VASP code [10, 11], but the number of basis functions needed is reduced.

II. METHOD OF CALCULATION

A. B-splines

B-splines form a convenient basis set with good flexibility. Here we briefly summarize the B-spline formalism. More details can be found in Ref. [7]. In general, B-spline of order κ consists of positive polynomials of degree $\kappa - 1$, over κ adjacent intervals. These polynomials vanish everywhere outside the subintervals $\tau_i < z < \tau_{i+\kappa}$. The B-spline basis set given by

the order κ and the knot sequence $\{\tau_i\}$ is generated by the following relation :

$$B_{i,\kappa}(z) = \frac{z - \tau_i}{\tau_{i+\kappa-1} - \tau_i} B_{i,\kappa-1}(z) + \frac{\tau_{i+\kappa} - z}{\tau_{i+\kappa} - \tau_{i+1}} B_{i+1,\kappa-1}(z), \quad (1)$$

with

$$B_{i,1}(z) = \begin{cases} 1, & \tau_i \leq z < \tau_{i+1} \\ 0, & \text{otherwise} . \end{cases} \quad (2)$$

The first derivative of the B-spline of order κ is given by

$$\frac{d}{dz} B_{i,\kappa}(z) = \frac{\kappa - 1}{\tau_{i+\kappa-1} - \tau_i} B_{i,\kappa-1}(z) - \frac{\kappa - 1}{\tau_{i+\kappa} - \tau_{i+1}} B_{i+1,\kappa-1}(z). \quad (3)$$

Since all the lower order B-splines are the simple polynomials, it is obvious that B-splines are superior to traditional methods in the treatment of derivatives. Figure 2 shows the B-splines generated from an equal distance and exponential type knot sequence. Clearly, B-splines possess good flexibility to accurately represent any localized function of z by suitably increasing the density of the knot sequence where it varies rapidly.

B. Relevant matrix elements within B-spline basis

1. NCPP scheme

In the NCPP scheme, the minimization of total energy subject to the constraint that the wave functions ϕ_i are orthogonal yields

$$H|\phi_i\rangle = \epsilon_i|\phi_i\rangle \quad (4)$$

where

$$H = -\nabla^2 + V_{pp} + V_H + V_{xc} . \quad (5)$$

Here, V_{pp} , V_H , and V_{xc} denote the pseudopotential, Hartree potential, and exchange-correlation potential, respectively.

The present mixed basis used to expand ϕ_i is defined as

$$\langle \mathbf{r} | \mathbf{k}_{\parallel} + \mathbf{G}_{\parallel}; j, \kappa \rangle = \frac{1}{\sqrt{A}} e^{i(\mathbf{k}_{\parallel} + \mathbf{G}_{\parallel}) \cdot \mathbf{r}_{\parallel}} B_{j,\kappa}(z) \quad (6)$$

where \mathbf{G}_{\parallel} denotes an in-plane reciprocal lattice vector, and $\mathbf{r}_{\parallel} = (\mathbf{x}, \mathbf{y})$ is the projection of \mathbf{r} in the x-y plane. \mathbf{k}_{\parallel} is the in-plane wave vector. A is the surface area of the system.

The overlap matrix elements between two basis states are given by

$$\langle \mathbf{k}_{\parallel} + \mathbf{G}_{\parallel}; i, \kappa \mid \mathbf{k}_{\parallel} + \mathbf{G}'_{\parallel}; i', \kappa \rangle = \langle B_{i,\kappa} \mid B_{i',\kappa} \rangle \delta_{\mathbf{G}_{\parallel}, \mathbf{G}'_{\parallel}} \quad (7)$$

where

$$\langle B_{i,\kappa} \mid B_{i',\kappa} \rangle = \int dz B_{i,\kappa}(z) B_{i',\kappa}(z) \quad (8)$$

is an integration of local polynomials with bounded support and vanishes unless the condition $|i - i'| \leq \kappa$ is fulfilled.

The kinetic energy matrix elements are given by

$$\begin{aligned} & \langle \mathbf{k}_{\parallel} + \mathbf{G}'_{\parallel}; i, \kappa \mid -\nabla^2 \mid \mathbf{k}_{\parallel} + \mathbf{G}_{\parallel}; i', \kappa \rangle \\ &= [\langle B_{i,\kappa} \mid -\frac{\partial^2}{\partial z^2} \mid B_{i',\kappa} \rangle + \langle B_{i,\kappa} \mid B_{i',\kappa} \rangle (\mathbf{k}_{\parallel} + \mathbf{G}_{\parallel})^2] \delta_{\mathbf{G}_{\parallel}, \mathbf{G}'_{\parallel}} \\ &= [\langle B'_{i,\kappa} \mid B'_{i',\kappa} \rangle + \langle B_{i,\kappa} \mid B_{i',\kappa} \rangle (\mathbf{k}_{\parallel} + \mathbf{G}_{\parallel})^2] \delta_{\mathbf{G}_{\parallel}, \mathbf{G}'_{\parallel}} . \end{aligned} \quad (9)$$

$B'_{i,\kappa}(z)$, the derivatives of $B_{i,\kappa}(z)$, can be expressed as a linear combination of $B_{i,\kappa-1}(z)$ with order $\kappa - 1$ and is continuous across the knot sequence.

The local part of V_{pp} on each atomic site with species σ concerned here can be written as

$$V_{\text{loc}}^{\sigma}(\mathbf{r}) = -\frac{Z^{\sigma}}{r} \text{erf}\left(\frac{r}{R_c^{\sigma}}\right) + \sum_i A_i^{\sigma} e^{-a_i^{\sigma} r^2}. \quad (10)$$

The first term on the right hand side of the above equation will be referred to as the core term due to the core charge distribution

$$\rho_c(r) = \frac{Z^{\sigma}}{\pi^{\frac{3}{2}} R_c^{\sigma 3}} e^{-\frac{r^2}{R_c^{\sigma 2}}}.$$

The local pseudopotential of the crystal is then given by

$$V_{\text{LOC}}(\mathbf{r}) = \sum_{\sigma, \mathbf{R}^{\sigma}} V_{\text{loc}}^{\sigma}(\mathbf{r} - \mathbf{R}^{\sigma}), \quad (11)$$

where \mathbf{R}^{σ} denotes the position of each atom with species σ . The matrix elements for the local pseudopotential of the crystal excluding the core term, V'_{LOC} , are given by

$$\begin{aligned} & \langle \mathbf{k}_{\parallel} + \mathbf{G}_{\parallel}; i, \kappa \mid V'_{\text{LOC}} \mid \mathbf{k}_{\parallel} + \mathbf{G}'_{\parallel}; i', \kappa \rangle = \\ & \sum_{\sigma, R_z \in \text{SUC}} I(i, \kappa; a_i^{\sigma}, R_z^{\sigma}; i', \kappa) e^{-i\Delta \mathbf{G}_{\parallel} \cdot \mathbf{R}_{\parallel}^{\sigma}} \tilde{V}_{\text{loc}}^{\sigma}(\Delta \mathbf{G}_{\parallel}) \end{aligned} \quad (12)$$

where $\Delta \mathbf{G}_{\parallel} = \mathbf{G}_{\parallel} - \mathbf{G}'_{\parallel}$ and

$$\begin{aligned} \tilde{V}_{\text{loc}}^{\sigma}(\Delta \mathbf{G}_{\parallel}) &= \iint d^2 \mathbf{r}_{\parallel} e^{-i(\Delta \mathbf{G}_{\parallel} \cdot \mathbf{r}_{\parallel} - a_i^{\sigma} r_{\parallel}^2)} \\ &= \sum_i \left(\frac{A_i^{\sigma}}{A_c} \right) \frac{\pi}{a_i^{\sigma}} e^{-\frac{\Delta \mathbf{G}_{\parallel}^2}{4a_i^{\sigma}}} . \end{aligned} \quad (13)$$

A_c is the area of the surface unit cell (SUC). I is given by

$$I(i, \kappa; a_i^{\sigma}, R_z^{\sigma}; i', \kappa) = \int dz B_{i, \kappa}(z) B_{i', \kappa}(z) e^{-a_i^{\sigma}(z - R_z^{\sigma})^2} .$$

The atomic nonlocal pseudopotential associated with species σ used in the present work is in the Kleinman-Bylander form [12], viz.,

$$V_{\text{nl}}^{\sigma}(\mathbf{r}) = \sum_{nlm} E_{nl}^{\sigma} |\beta_{n,lm}^{\sigma} \rangle \langle \beta_{n,lm}^{\sigma} | . \quad (14)$$

The projector $\beta_{n,lm}^{\sigma}$ is a radial function multiplied by an angular momentum eigenfunction $Y_{lm}(\mathbf{\Omega})$. For the norm-conserving pseudopotential by Goedecker, Teter, and Hutter (GTH) [8], there are two s -channels and one p -channel for the second row element,

$$\begin{aligned} \beta_{1,00}^{\sigma}(\mathbf{r}) &= C_{1,0}^{\sigma} e^{-\alpha_0^{\sigma} r^2} r^0 Y_{00}(\mathbf{\Omega}) \\ \beta_{2,00}^{\sigma}(\mathbf{r}) &= C_{2,0}^{\sigma} e^{-\alpha_0^{\sigma} r^2} r^2 Y_{00}(\mathbf{\Omega}) \\ \beta_{1,1m}^{\sigma}(\mathbf{r}) &= C_{1,1}^{\sigma} e^{-\alpha_1^{\sigma} r^2} r^1 Y_{1m}(\mathbf{\Omega}) . \end{aligned} \quad (15)$$

Since the d -like pseudopotential is chosen as the local pseudopotential, only $l = 0$ and $l = 1$ are considered here. The nonlocal pseudopotential of the crystal is

$$V_{\text{NL}}(\mathbf{r}) = \sum_{\sigma, \mathbf{R}^{\sigma}} V_{\text{nl}}^{\sigma}(\mathbf{r} - \mathbf{R}^{\sigma}) \quad (16)$$

and the matrix elements of V_{NL} are given by

$$\begin{aligned} &\langle \mathbf{k}_{\parallel} + \mathbf{G}_{\parallel}; i, \kappa \mid V_{\text{NL}} \mid \mathbf{k}_{\parallel} + \mathbf{G}'_{\parallel}; i', \kappa \rangle \\ &= \sum_{\sigma, \mathbf{R}^{\sigma}} \sum_{nlm} \langle \mathbf{k}_{\parallel} + \mathbf{G}_{\parallel}; i, \kappa \mid \beta_{n,lm}^{\sigma, \mathbf{R}^{\sigma}} \rangle E_{nl}^{\sigma} \langle \beta_{n,lm}^{\sigma, \mathbf{R}^{\sigma}} \mid \mathbf{k}_{\parallel} + \mathbf{G}'_{\parallel}; i', \kappa \rangle \end{aligned} \quad (17)$$

where $\beta_{n,lm}^{\sigma, \mathbf{R}^{\sigma}} \equiv \beta_{n,lm}^{\sigma}(\mathbf{r} - \mathbf{R}^{\sigma})$. By the representation of $r^l Y_{lm}(\mathbf{\Omega})$ in Cartesian coordinates for $l = 0$ and $l = 1$,

$$\begin{aligned} r^0 Y_{00}(\mathbf{\Omega}) &= \frac{1}{\sqrt{4\pi}}, & r^1 Y_{1-1}(\mathbf{\Omega}) &= \sqrt{\frac{3}{8\pi}} (x - iy) \\ r^1 Y_{10}(\mathbf{\Omega}) &= \sqrt{\frac{3}{4\pi}} z, & r^1 Y_{11}(\mathbf{\Omega}) &= -\sqrt{\frac{3}{8\pi}} (x + iy) \\ r^2 Y_{00}(\mathbf{\Omega}) &= \frac{1}{\sqrt{4\pi}} (x^2 + y^2 + z^2) \end{aligned}$$

we obtain

$$\begin{aligned}
& \langle \mathbf{k}_{\parallel} + \mathbf{G}_{\parallel}; i, \kappa | V_{\text{NL}} | \mathbf{k}_{\parallel} + \mathbf{G}'_{\parallel}; i', \kappa \rangle = \\
& \frac{1}{A_c} \sum_{\sigma, \mathbf{R}^{\sigma} \in \text{SUC}} \sum_{nlm} e^{-i\Delta\mathbf{G}_{\parallel} \cdot \mathbf{R}_{\parallel}^{\sigma}} E_{nl}^{\sigma} [C_{n,l}^{\sigma} S_{n,lm}^{\sigma}(i, \kappa; \mathbf{k}_{\parallel} + \mathbf{G}_{\parallel}; R_z^{\sigma}, \alpha_l^{\sigma})]^* \\
& \times [C_{n,l}^{\sigma} S_{n,lm}^{\sigma}(i', \kappa; \mathbf{k}_{\parallel} + \mathbf{G}'_{\parallel}; R_z^{\sigma}, \alpha_l^{\sigma})]
\end{aligned}$$

where

$$S_{n,lm}^{\sigma}(i, \kappa; \mathbf{k}_{\parallel} + \mathbf{G}_{\parallel}; R_z, \alpha_l^{\sigma}) = \begin{cases} Z_{1,lm} Q_{1,lm}, & n = 1 \\ Z_{1,00} Q_{2,00} + Z_{2,00} Q_{1,00}, & n = 2 \end{cases}$$

with the factors $Z_{n,lm}$ and $Q_{n,lm}$ given explicitly as

$$\begin{aligned}
Z_{1,00} &= \int dz B_{i,\kappa}(z) e^{-\alpha_0^{\sigma}(z-R_z^{\sigma})^2} \\
Z_{1,1\pm 1} &= \int dz B_{i,\kappa}(z) e^{-\alpha_1^{\sigma}(z-R_z^{\sigma})^2} \\
Z_{1,10} &= \int dz B_{i,\kappa}(z) (z - R_z^{\sigma}) e^{-\alpha_1^{\sigma}(z-R_z^{\sigma})^2} \\
Z_{2,00} &= \int dz B_{i,\kappa}(z) (z - R_z^{\sigma})^2 e^{-\alpha_0^{\sigma}(z-R_z^{\sigma})^2} \\
Q_{1,10} &= \iint d^2\mathbf{r}_{\parallel} e^{i(\mathbf{k}_{\parallel} + \mathbf{G}_{\parallel}) \cdot \mathbf{r}_{\parallel}} e^{-\alpha_l^{\sigma} \mathbf{r}_{\parallel}^2} \\
&= \frac{\pi}{\alpha_l^{\sigma}} e^{-\frac{|\mathbf{k}_{\parallel} + \mathbf{G}_{\parallel}|^2}{4\alpha_l^{\sigma}}} \\
Q_{1,1\pm 1} &= - \iint d^2\mathbf{r}_{\parallel} \left(\frac{x \pm iy}{\sqrt{2}} \right) e^{i(\mathbf{k}_{\parallel} + \mathbf{G}_{\parallel}) \cdot \mathbf{r}_{\parallel}} e^{-\alpha_1^{\sigma} \mathbf{r}_{\parallel}^2} \\
&= - \left(\frac{1}{\sqrt{2}} \frac{k_x + G_x}{2\alpha_1^{\sigma}} \pm \frac{i}{\sqrt{2}} \frac{k_y + G_y}{2\alpha_1^{\sigma}} \right) \frac{\pi}{\alpha_1^{\sigma}} e^{-\frac{|\mathbf{k}_{\parallel} + \mathbf{G}_{\parallel}|^2}{4\alpha_1^{\sigma}}} \\
Q_{2,00} &= \iint d^2\mathbf{r}_{\parallel} (x^2 + y^2) e^{i(\mathbf{k}_{\parallel} + \mathbf{G}_{\parallel}) \cdot \mathbf{r}_{\parallel}} e^{-\alpha_0^{\sigma} \mathbf{r}_{\parallel}^2} \\
&= \left(\frac{1}{\alpha_0^{\sigma}} - \frac{|\mathbf{k}_{\parallel} + \mathbf{G}_{\parallel}|^2}{4\alpha_0^{\sigma 2}} \right) \frac{\pi}{\alpha_0^{\sigma}} e^{-\frac{|\mathbf{k}_{\parallel} + \mathbf{G}_{\parallel}|^2}{4\alpha_0^{\sigma}}} .
\end{aligned}$$

The total charge distribution ρ is defined as the sum of the core charge distributions for all atoms in the sample ρ_c , plus the electronic charge distributions ρ_e ,

$$\rho(\mathbf{r}) = \rho_c(\mathbf{r}) + \rho_e(\mathbf{r}) = \rho_c(\mathbf{r}) + \sum_i |\phi_i(\mathbf{r})|^2 . \quad (18)$$

The remainder of the crystal potential includes the Hartree potential V_H due to the electron charge distribution, the exchange-correlation potential V_{xc} , and the core term of the local

pseudopotential $V_{\text{LOC}}^{(C)}$ omitted in the above. Let $V^{scf} \equiv V_H + V_{xc} + V_{\text{LOC}}^{(C)}$. This potential is local and periodic in the plane. The exchange-correlation potential used here are deduced from the Monte Carlo results calculated by Ceperley and Alder[13] and parametrized by Perdew and Zunger[14]. We write

$$V^{scf}(\mathbf{r}) = \sum_{\mathbf{g}_{\parallel}} V^{scf}(z, \mathbf{g}_{\parallel}) e^{i\mathbf{g}_{\parallel} \cdot \mathbf{r}_{\parallel}} . \quad (19)$$

The matrix elements of V^{scf} are given by

$$\langle \mathbf{k}_{\parallel} + \mathbf{G}_{\parallel}; i, \kappa | V^{scf} | \mathbf{k}_{\parallel} + \mathbf{G}'_{\parallel}; i', \kappa \rangle = \int dz B_{i;\kappa}(z) B_{i';\kappa}(z) V^{scf}(z, \Delta \mathbf{G}_{\parallel}) . \quad (20)$$

With the use of the 2D Fourier transformation of $1/r$

$$\frac{1}{r} = 4\pi \iint \frac{d^2 \mathbf{k}_{\parallel}}{(2\pi)^2} e^{i\mathbf{k}_{\parallel} \cdot \mathbf{r}_{\parallel}} \frac{1}{2k_{\parallel}} e^{-k_{\parallel}|z|}$$

the Coulomb potential due to the total charge distribution is given by

$$\begin{aligned} V^{(C)} &= V_{\text{LOC}}^{(C)} + V_H = \iiint d^3 \mathbf{r}' \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \\ &= \sum_{\mathbf{g}_{\parallel}} \left(\int dz' \rho(z', \mathbf{g}_{\parallel}) \frac{2\pi}{|\mathbf{g}_{\parallel}|} e^{-|\mathbf{g}_{\parallel}| |z-z'|} \right) e^{-i\mathbf{g}_{\parallel} \cdot \mathbf{r}_{\parallel}} \\ &\equiv \sum_{\mathbf{g}_{\parallel}} V^{(C)}(z, \mathbf{g}_{\parallel}) e^{-i\mathbf{g}_{\parallel} \cdot \mathbf{r}_{\parallel}} \end{aligned} \quad (21)$$

where

$$\rho(z, \mathbf{g}_{\parallel}) = \frac{1}{A} \iint d^2 \mathbf{r}_{\parallel} \rho(\mathbf{r}) e^{i\mathbf{g}_{\parallel} \cdot \mathbf{r}_{\parallel}} .$$

2. USPP scheme

For Vanderbilt's ultra-soft pseudopotentials [9, 15], Eq. (4) becomes a secular equation of the form

$$H|\phi_i \rangle = \epsilon_i S|\phi_i \rangle \quad (22)$$

under a generalized orthonormality condition

$$\langle \phi_i | S | \phi_j \rangle = \delta_{ij} . \quad (23)$$

S is a Hermitian overlap operator given by

$$S = I + \sum_{\sigma, \mathbf{R}^{\sigma}} \sum_{nlmn'l'm'} q_{nlm, n'l'm'}^{\sigma} |\beta_{n, lm}^{\sigma, \mathbf{R}^{\sigma}} \rangle \langle \beta_{n', l'm'}^{\sigma, \mathbf{R}^{\sigma}}| , \quad (24)$$

where $q_{nlm,n'l'm'}^\sigma = \iiint d^3\mathbf{r} Q_{nlm,n'l'm'}^\sigma(\mathbf{r})$. Here, $\beta_{n,lm}^\sigma$ and $Q_{nlm,n'l'm'}^\sigma$ vanish outside the core region.

The individual atomic nonlocal potential in Eq. (16) is modified as

$$V_{nl}^\sigma(\mathbf{r} - \mathbf{R}^\sigma) = \sum_{nlmn'l'm'} E_{nlm,n'l'm'}^{\sigma,\mathbf{R}^\sigma} |\beta_{n,lm}^{\sigma,\mathbf{R}^\sigma} \rangle \langle \beta_{n',l'm'}^{\sigma,\mathbf{R}^\sigma}|, \quad (25)$$

with

$$E_{nlm,n'l'm'}^{\sigma,\mathbf{R}^\sigma} = E_{nl,n'l'}^{0,\sigma} \delta_{l,l'} \delta_{m,m'} + \iiint d^3\mathbf{r} V_{eff}(\mathbf{r}) Q_{nlm,n'l'm'}^\sigma(\mathbf{r} - \mathbf{R}^\sigma). \quad (26)$$

V_{eff} is defined as

$$V_{eff} = V_H + V_{xc} + V_{LOC} = V^{scf} + V'_{LOC}. \quad (27)$$

Following Ref. [15], we define a box, which is large enough to contain the core region. The USPP $\beta_{n,lm}^\sigma(z - R_z^\sigma, \mathbf{r}_\parallel)$ inside the box is transferred to \mathbf{G}_\parallel space using the fast Fourier transform (FFT), then

$$\beta_{n,lm}^{\sigma,\mathbf{R}^\sigma} = \beta_{n,lm}^\sigma(\mathbf{r} - \mathbf{R}^\sigma) = \sum_{\mathbf{G}_\parallel} \beta_{n,lm}^\sigma(z - R_z^\sigma, \mathbf{G}_\parallel) e^{-i\mathbf{G}_\parallel \cdot \mathbf{R}^\sigma} e^{i\mathbf{G}_\parallel \cdot \mathbf{r}_\parallel}. \quad (28)$$

So, the projection of $\beta_{n,lm}^\sigma(\mathbf{r} - \mathbf{R}^\sigma)$ on the basis in Eq. (17) reads as

$$\begin{aligned} & \langle \mathbf{k}_\parallel + \mathbf{G}_\parallel; i, \kappa | \beta_{n,lm}^{\sigma,\mathbf{R}^\sigma} \rangle \\ &= e^{-i(\mathbf{G}_\parallel + \mathbf{k}_\parallel) \cdot \mathbf{R}^\sigma} \sum_{\mathbf{G}'_\parallel} \int dz B_{i,\kappa}(z) \beta_{n,lm}^\sigma(z - R_z^\sigma, \mathbf{G}'_\parallel) \iint d^2\mathbf{r}_\parallel e^{i(\mathbf{G}'_\parallel - \mathbf{G}_\parallel - \mathbf{k}_\parallel) \cdot \mathbf{r}_\parallel}. \end{aligned} \quad (29)$$

Presently, we restrict ourselves to

$$Q_{nlm,n'l'm'}^\sigma(\mathbf{r}) = Q_{nl,n'l'}^\sigma(r) Y_{LM} \quad (30)$$

with L be the minimum l_{min} of $|l-l'|$ and $|M = m+m'| \leq l_{min}$. Similarly, both $Q_{nlm,n'l'm'}^\sigma(\mathbf{r} - \mathbf{R}^\sigma)$ and $V_{eff}(\mathbf{r})$ in Eq. (26) are also transferred by FFT,

$$Q_{nlm,n'l'm'}^\sigma(\mathbf{r} - \mathbf{R}^\sigma) = \sum_{\mathbf{g}_\parallel} Q_{nlm,n'l'm'}^\sigma(z - R_z^\sigma, \mathbf{R}_\parallel^\sigma; \mathbf{g}_\parallel) e^{i\mathbf{g}_\parallel \cdot \mathbf{r}_\parallel}, \quad (31)$$

$$V_{eff}(\mathbf{r}) = \sum_{\mathbf{g}_\parallel} V_{eff}(z, \mathbf{g}_\parallel) e^{i\mathbf{g}_\parallel \cdot \mathbf{r}_\parallel}. \quad (32)$$

Then, we obtain

$$E_{nlm,n'l'm'}^{\sigma,\mathbf{R}^\sigma} = E_{nl,n'l'}^{0,\sigma} \delta_{l,l'} \delta_{m,m'} + A_b^\sigma \sum_{\mathbf{g}_\parallel} \int dz [V_{eff}(z, \mathbf{g}_\parallel)]^* Q_{nlm,n'l'm'}^\sigma(z - R_z^\sigma, \mathbf{R}_\parallel^\sigma; \mathbf{g}_\parallel). \quad (33)$$

A_b^σ is the surface area of the core region box. Note that the FFT grid density for \mathbf{g}_\parallel in the summation of Eqs. (31) and (32) is not necessarily the same with that for the wavefunction [15]. $E_{nlm,n'l'm'}^{\sigma,\mathbf{R}^\sigma}$ in the above should be calculated self-consistently. From Eqs. (29) and (33), we can evaluate $V_{\text{NL}}|\phi_i\rangle$ and $S|\phi_i\rangle$.

Finally, the charge density from the wave function is augmented inside the core region,

$$\rho_e(\mathbf{r}) = \sum_i [|\phi_i(\mathbf{r})|^2 + \sum_{\sigma,\mathbf{R}^\sigma} \sum_{nlmn'l'm'} Q_{nlm,n'l'm'}^\sigma(\mathbf{r} - \mathbf{R}^\sigma) \langle \phi_i | \beta_{n,lm}^{\sigma,\mathbf{R}^\sigma} \rangle \langle \beta_{n',l'm'}^{\sigma,\mathbf{R}^\sigma} | \phi_i \rangle]. \quad (34)$$

These formula can be easily extended for the 1D case, i.e., using two sets of B-splines to describe the non-periodic directions and 1D plane waves for the periodic one.

C. Generalized Lanczos-Krylov method for diagonalization

In most cases, the only practical approach to find the lowest eigenvectors of the Hamiltonian matrix H with a very large dimension is through iterations. In particular, it is known [16] that the use of Krylov subspaces provides stability for the iteration process.

We start with a diagonalization in the subspace of the union of all the input trial vectors with the Lanczos process [17] and the Krylov subspace K generated by repeated operations of H on one of the trial vectors \mathbf{v} :

$$K = \text{Span}\{\mathbf{v}, H\mathbf{v}, H^2\mathbf{v}, \dots, H^k\mathbf{v}\}.$$

The orthonormality of the vectors is maintained throughout by the standard Gram-Schmidt orthogonalization procedure. Thus a set of normalized roots arranged according to their eigenvalues are obtained.

In the next step, these roots are divided into blocks of small sets of the trial vectors. In each block, the diagonalization is then performed in the subspace of these vectors together with a Krylov subspace K based on one of the trial vectors in the block. Again the orthonormality of the vectors is maintained throughout by the Gram-Schmidt scheme. Maximum overlapping with the input trial vectors in each block is used as the criteria for the selection of the desired roots. This criterion guarantees the stability of the iteration process while allowing the diagonalization to be carried out on subsets with a small number of vectors. Additional procedures are inserted to pick up possible additional new roots in each block. Thus, a set of improved normalized roots in each block are obtained.

Due to its stability, the Jacobi method is used for the diagonalization in each step. Modern Jacobi methods can compete in speed with Householder-based algorithms. Also, it can be easily parallelized. An additional advantage of the Jacobi method is that it takes full advantage of the progressive iterative approximation property of the input eigenvectors. This makes it especially suitable for iteration procedures.

These two alternating grand cycles are then iterated. It is found that only a few iterations are needed.

III. APPLICATIONS OF PRESENT METHOD

To demonstrate the capability of the present method, we apply it to study a few examples, including the C(001)-(2×1) surface, N δ -doped graphene, and the graphene nanoribbon.

A. C(001)-(2×1) surface

We first calculated the band structure of the C(001)-(2×1) surface simulated by a ten-layer slab. The relaxed positions of the surface layer are taken from Ref. [18], as displayed in Fig. 3(a). We did not try to determine the optimum geometry presently, which will be implemented in the future. A total of 40 B-splines, defined over a range of $4.5 a_0$ ($a_0 = 3.52 \text{ \AA}$), are used to expand the z -component wavefunction. The energy cutoff E_c of the 2D plane waves is 50 Ry. The 3×6 Monkhorst-Pack grids were taken to sample the surface Brillouin zone. We used the GTH norm-conserving pseudopotential [8] for the interactions between the ions and valence electrons. The potential V^{scf} is determined self-consistently until its change is less than 10^{-6} Ry. For comparison, we also performed the calculation by using the VASP code with the projector-augmented-wave method (PAW) [10, 11].

In practical calculations, the significant charge-density oscillation, the so-called charge sloshing, was observed. To remedy the very slow convergence by the simple linear mixing scheme (Fig. 4), we use the Kerker mixing [19] to prevent the charge sloshing and the residual minimization method in the direct inversion of iterative subspace (RMM-DIIS) Pulay scheme [20, 21] to accelerate the convergence. The new charge is mixed by

$$\rho_{new}(z, \mathbf{g}_{\parallel}) = \rho_{in}(z, \mathbf{g}_{\parallel}) + W \frac{\mathbf{g}_{\parallel}^2}{\mathbf{g}_{\parallel}^2 + \mathbf{g}_{\parallel, \max}^2} (\rho_{out}(z, \mathbf{g}_{\parallel}) - \rho_{in}(z, \mathbf{g}_{\parallel}))$$

The weight factor W and the cutoff wavevector $\mathbf{g}_{\parallel, \text{max}}$ are two adjustable parameters. The updated input charge density, as a linear combination of charge density of all 4 previous steps, is determined such that it minimizes the corresponding residual vector.

Figures 3(b) and (c) display the results. The bands for wave vectors along $X' - M - X$ should be twofold degenerate due to the symmetry of the slab system [22]. In the present calculation, the corresponding splitting is $< 10^{-5}$ eV for all \mathbf{k} along $X' - M - X$. As can be seen in Fig. 3(b) and (c), our results are almost identical to those by the VASP code. The filled π band and the empty π^* band, separated in energy by ~ 2 eV, are clearly seen in the bulk band gap, and are also in nice agreement with those in Ref. [23].

B. n-type δ -doped graphene

Before proceeding to the second example, we should keep in mind that, although the utility of first-principles norm-conserving pseudopotentials has paved the way to accurate calculations of solid-state properties, the norm-conserving constraint is the main factor responsible for the hardness of highly localized valence orbitals of $1s$, $2p$, or $3d$, with no core state of the same angular momentum. Therefore, the all-electron wave function is nodeless and quite compressed compared to the other valence states, thus requiring a large number of plane waves to be represented accurately. For the next two graphene-related examples, we will use the Vanderbilt's USPP scheme in which such a constraint is relaxed. In this scheme, the pseudo-wave functions are allowed to be as soft as possible within a given region, yielding a dramatic reduction of the cutoff energy. For instance, we found that the change in the band structure of the graphene by the NCPP will be unnoticeable with $E_c = 70$ Ry but, by the USPP with $E_c = 20$ Ry only. This reduction of the cutoff energy in USPP is particularly important for large systems to avoid the diagonalization bottleneck of the extremely large matrix H .

First of all, we examined the quality of the C and N USPPs that were generated from the Vanderbilt's code [24], by performing the band structure calculation for the graphene and the assumed N-doped graphene with equal number of C and N atoms. Figure 5 displays the band structures, along with the VASP-PAW results. Clearly, both are nearly identical for these two cases. Therefore, we believe that the quality of C and N USPPs used here is sufficiently good for our band structure calculations.

The calculations were carried out using a supercell of $8 \times \sqrt{3}$, i.e., 16 carbon dimmer lines between adjacent N chains, where the N chain is the δ dopant. All C and N atoms were kept at the ideal positions, with the lattice constant a_0 set to be 2.641 Å. A mixed basis set with 13 B-splines distributed over a range of $3.25 a_0$ and the plane wave cutoff of 20 Ry are used. The 2×4 Monkhorst-Pack grids were taken to sample the surface Brillouin zone.

Overall, we found an excellent agreement between the present band structures and the VASP-PAW results. Figure 6 shows only the results near the Fermi level along $\Gamma - X$. We have also tested various N mole fractions x of 1/2, 1/4, and 1/8, with x defined as the ratio of the number of N to the total number of atoms per unit cell. Besides all agreeable with the VASP-PAW results, we found that the trend of the band gap is also in accord with previous work [25], e.g., the band gap opens only for large x (Fig. 5(c)). Note that the USPP and PAW methods have proven to be closed related [10]. Therefore, we are confident that the present program has produced reliable results.

It is worth noting that the ratio of the number of the basis N_d used in VASP and the present approach is about 1.4 – 1.7 for $E_c = 20 - 30$ Ry, with a typical thickness of 10 Å for the vacuum layer required in VASP. Like in the conjugate gradient technique [11, 26, 27], the computational effort by the present Lanczos-Krylov method scales as $M^2 N_d$ for the orthogonalization and $M N_d \log N_d$ for the FFT, where M is the number of the bands considered. Therefore, as compared to the conventional supercell modeling, the computational time needed for the eigenvalues search due to the planar mixed basis will reduce with a factor of 2 - 3 if the same algorithm were used. This reduction will be more significant for the cases, in which the tail of the Coulomb potential extends to a very long distance, whereas the wave function still decays exponentially in the vacuum.

C. graphene nanoribbon

Finally, we will apply the present method to armchair graphene nanoribbon. In the spirit of our method, we should use two sets of B-splines describing the two non-periodic dimensions and only 1D plane waves for the periodic y direction in such a system. Here, we just adopt the planar mixed basis set to study the nanoribbon by using the surface supercell modeling for the $x - y$ plane.

Armchair graphene nanoribbons are characterized by the number of dimmer lines N_a

(Fig. 7(a)) across the ribbons. The width of the ribbon chosen is chosen to include $N_a = 19$ carbon dimer lines, which is $3p + 1$ type (p is a integer). We performed the calculation using a $16 \times \sqrt{3}$ surface supercell with the vacuum 'band' of ~ 15 Å along the y direction. The optimized atomic positions in the eight outer dimer lines of both sides obtained by the VASP code were used for the calculation. Since graphene nanoribbons are stripes of graphene, edge atoms are not saturated [28]. For simplicity, we let the edge atoms not saturated by hydrogen atoms in the present calculation. All other computational conditions are similar to the case of n-type δ -doped graphene described above.

The band structures of the present USPP result and the VASP-PAW counterpart near the Fermi level are displayed in Fig. 7(b) and (c), respectively. Again, there is a very nice agreement between these two approaches. It has been reported [29] that all armchair graphene nanoribbons with the three typical families, i.e., $N_a = 3p$, $3p + 1$, and $3p + 2$, are semiconductors with a direct energy gap. The bandgap obtained by our approach is 0.57 eV, and 0.61 eV by the VASP code. These values compare favorably with that in Ref. [29].

To sum up, we are convinced that the present program has been implemented successfully for low-dimensional systems and the results obtained are very reliable. Moreover, as compared to traditional supercell modeling, the use of the present mixed basis will reduce the number of the basis functions and speed up the calculations of electronic states.

IV. CONCLUSIONS

In conclusion, we have developed an efficient and accurate method to investigate the electronic structures of low dimensional systems by a mixed basis set with plane waves for the periodic direction and B-spline polynomials for the non-periodic direction. Contrary to the existing algorithms based upon the conventional supercell model with alternating slab and vacuum regions, it is a real space approach along the z -axis and therefore gives the surface band structures in absolute-energy scale. Furthermore, B-splines are independent of atomic positions when the atomic structure is optimized; thus, the geometry optimization can be easily implemented. We have also implemented a generalized Lanczos-Krylov iterative method for the diagonalization of the Kohn-Sham Hamiltonian. It offers a promising way to find the lower eigenvalues of a real matrix with a very large dimension through iterations. Particularly, this algorithm is very stable for low dimensional systems, of which the wave

functions vary abruptly.

We have calculated the electronic structures of the C(001)-(2×1) surface with GTH norm-conserving pseudopotentials [8] and the N δ -doped graphene, and graphene nanoribbon with Vanderbilt's ultra-soft pseudopotentials [9, 15]. It is found that the band structures are all in good agreement with those by the popular existing codes, but with a reduced number of basis functions.

Acknowledgments

This work was supported by the National Science Council under grant numbers NSC 100-2112-M-017-002-MY3 and NSC 01-2112-M-001-024-MY3 and by National Center for Theoretical Sciences of Taiwan.

-
- [1] G.-W. Li and Y.-C. Chang, Phys. Rev. B **48**, 12032 (1993).
 - [2] G.-W. Li and Y.-C. Chang, Phys. Rev. B **50**, 8675 (1994).
 - [3] Y.-C. Chang and G.-W. Li, Comp. Phys. Comm. **95**, 158 (1996).
 - [4] W. R. Johnson, S. A. Blundell, and J. Sapirstein, Phys. Rev. A **37**, 307 (1988).
 - [5] H. T. Jeng, and C. S. Hsue, Phys. Rev. B **62**, 9876 (2000).
 - [6] C. Y. Ren, H. T. Jeng, and C. S. Hsue, Phys. Rev. B **66**, 125105 (2002).
 - [7] Carl deBoor, *A practical Guide to Splines*, (Springer, New York, 1987).
 - [8] S. Goedecker, M. Teter, and J. Hutter, Phys. Rev. B **54**, 1703 (1996).
 - [9] D. Vanderbilt, Phys. Rev. B **41**, 7982 (1990).
 - [10] G. Kresse and D. Joubert, Phys. Rev. B **59**, 1758 (1999).
 - [11] G. Kresse and J. Furthmüller, Comput. Mater. Sci. **6**, 15 (1996).
 - [12] L. Kleinman and D. M. Bylander, Phys. Rev. Lett. **48**, 1425 (1982).
 - [13] D. M. Ceperley and B. J. Alder, Phys. Rev. Lett. **45**, 566 (1980).
 - [14] J. P. Perdew and A. Zunger, Phys. Rev. B **23**, 5048 (1981).
 - [15] K. Laasonen, A. Pasquarello, R. Car, C. Lee, and D. Vanderbilt, Phys. Rev. B **47**, 10142 (1993).
 - [16] For example, J. Dongarra and F. Sullivan, Comput. Sci. Eng., **2**, 22 (2000).

- [17] C. Lanczos, J. Res. Nat. Bur. Stand., **45**, 255 (1950).
- [18] Z. Jing and J. L. Whitten, Phys. Rev. B **50**, 2598 (1994).
- [19] G. P. Kerker, Phys. Rev. B **23**, 3062 (1981).
- [20] P. Csaszar and P. Pulay, J. Mol. Struct. **114**, 31 (1984).
- [21] F. Eckert, P. Pulay, and H.-J. Werner, J. Comp. Chem. **18**, 1473 (1997).
- [22] Here, we use the symmetric dimer model for C(001)-(2×1) reconstrued surface for testing the present approach.
- [23] P. Kröger and J. Pollmann, Phys. Rev. Lett. **74**, 1155 (1995).
- [24] <http://www.physics.rutgers.edu/~dhv/uspp/>.
- [25] X.-L. Wei, H. Fang, R. -Z. Wang, Y.-P. Chen and J.-X. Zhong , Appl. Phys. Lett. **99**, 012107 (2011).
- [26] M. P. Teter, M. C. Payne, and D. C. Allan, Phys. Rev. B **40**, 12255 (1989).
- [27] M. C. Payne, M. P. Teter and D. C. Allan, Rev. Mod. Phys. **64**, 1045 (1992).
- [28] E. Kan, Z.-Y. Li and J.-L. Yang, *Graphene Nanoribbons: Geometric, Electronic, and Magnetic Properties* in *Physics and Applications of Graphene - Theory* edited by S. Mikhailov, (Intech, 2011).
- [29] Y.-W. Son, M. L. Cohen, and S. G. Louie, Phys. Rev. Lett. **97**, 216803, (2006).

FIGURE CAPTIONS

Fig. 1: Schematic plot of a real slab, investigated by (a) the conventional supercell model and by (b) the real space approach with a localized basis set.

Fig. 2: B-splines defined on (a) an equal-distance (b) an exponential-type knot sequence (noted as $*$) of order $\kappa = 5$.

Fig. 3: (Color online) (a) Atomic structure of the C(001)-(2 \times 1) surface. Atoms located in deeper layer are denoted by smaller circles. (b) and (c) are the corresponding band structures obtained by the present work with NCPP and VASP with PAW.

Fig. 4: (Color online) Convergence rate for the calculation of the C(001)-(2 \times 1) surface by the Pulay-Kerker mixing and simple linear mixing scheme.

Fig. 5: Band structures of graphene obtained by (a) the present work with USPP and (b) VASP with PAW. (c) and (d) are the band structures of N δ -doped graphene obtained by the present work with USPP and VASP with PAW, respectively.

Fig. 6: (Color online) (a) Atomic structure of the N δ -doped graphene. Red circles denote N atoms. (b) and (c) are the corresponding band structures by the present work with USPP and VASP with PAW.

Fig. 7: (Color online) (a) Atomic structure of the armchair graphene nanoribbon. (b) and (c) are the corresponding band structures by the present work with USPP and VASP with PAW.

FIG. 1:

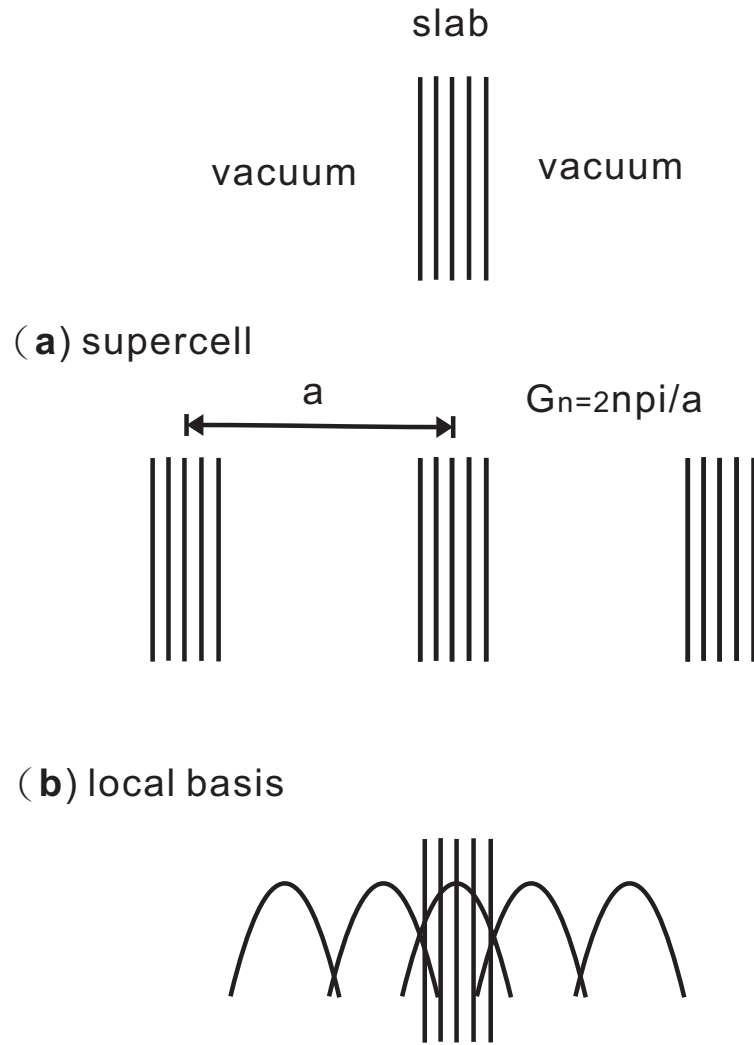


FIG. 2:

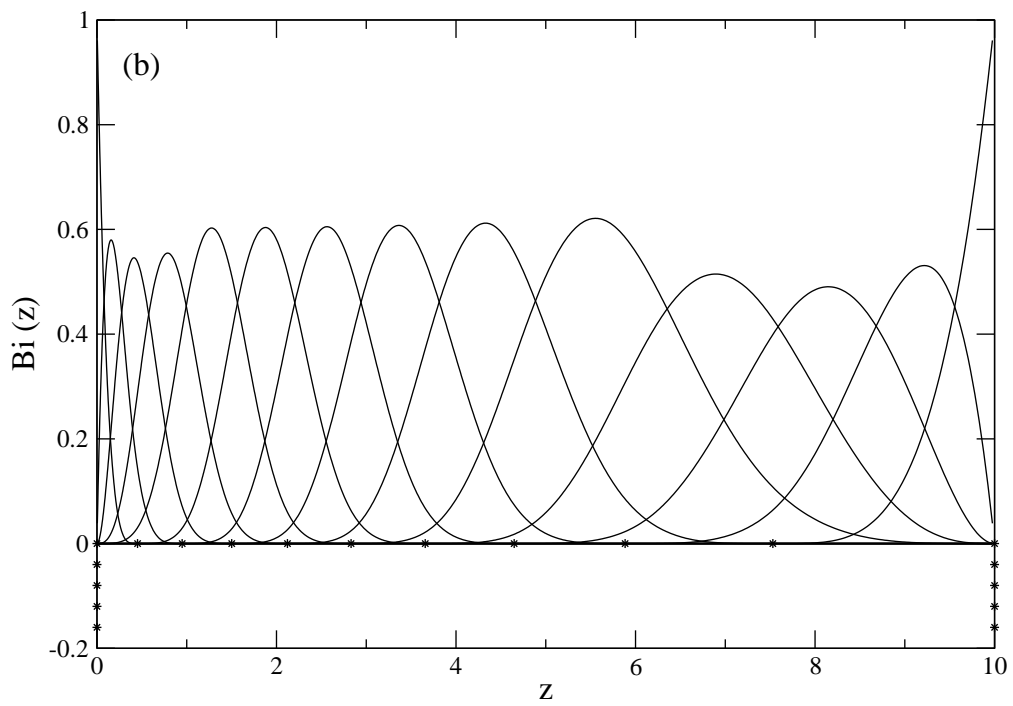
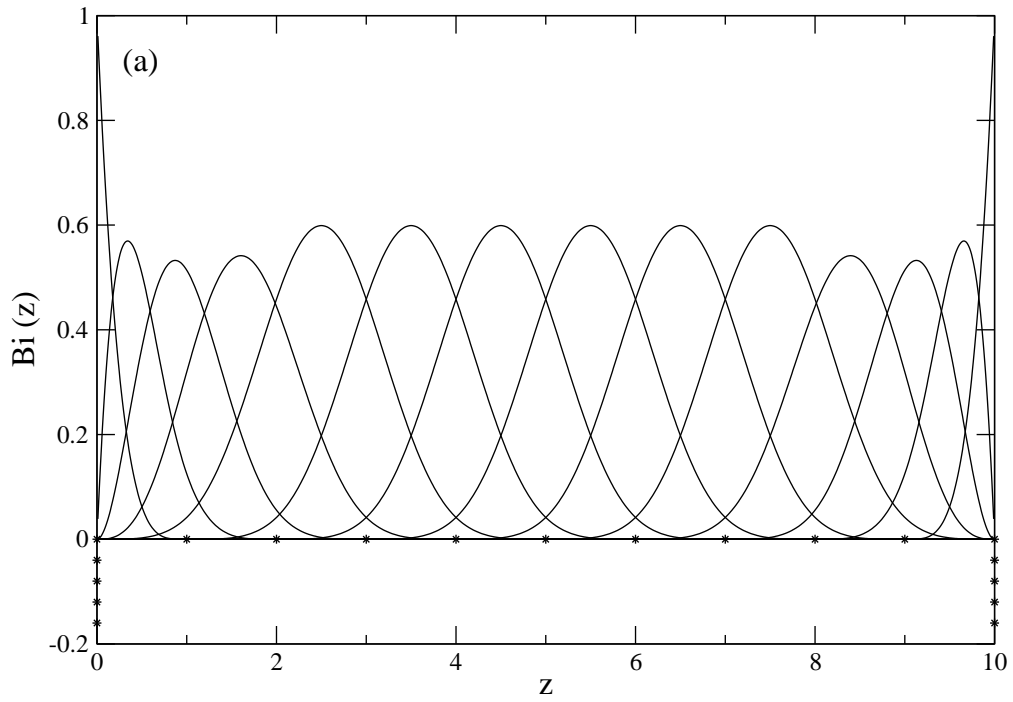


FIG. 3:

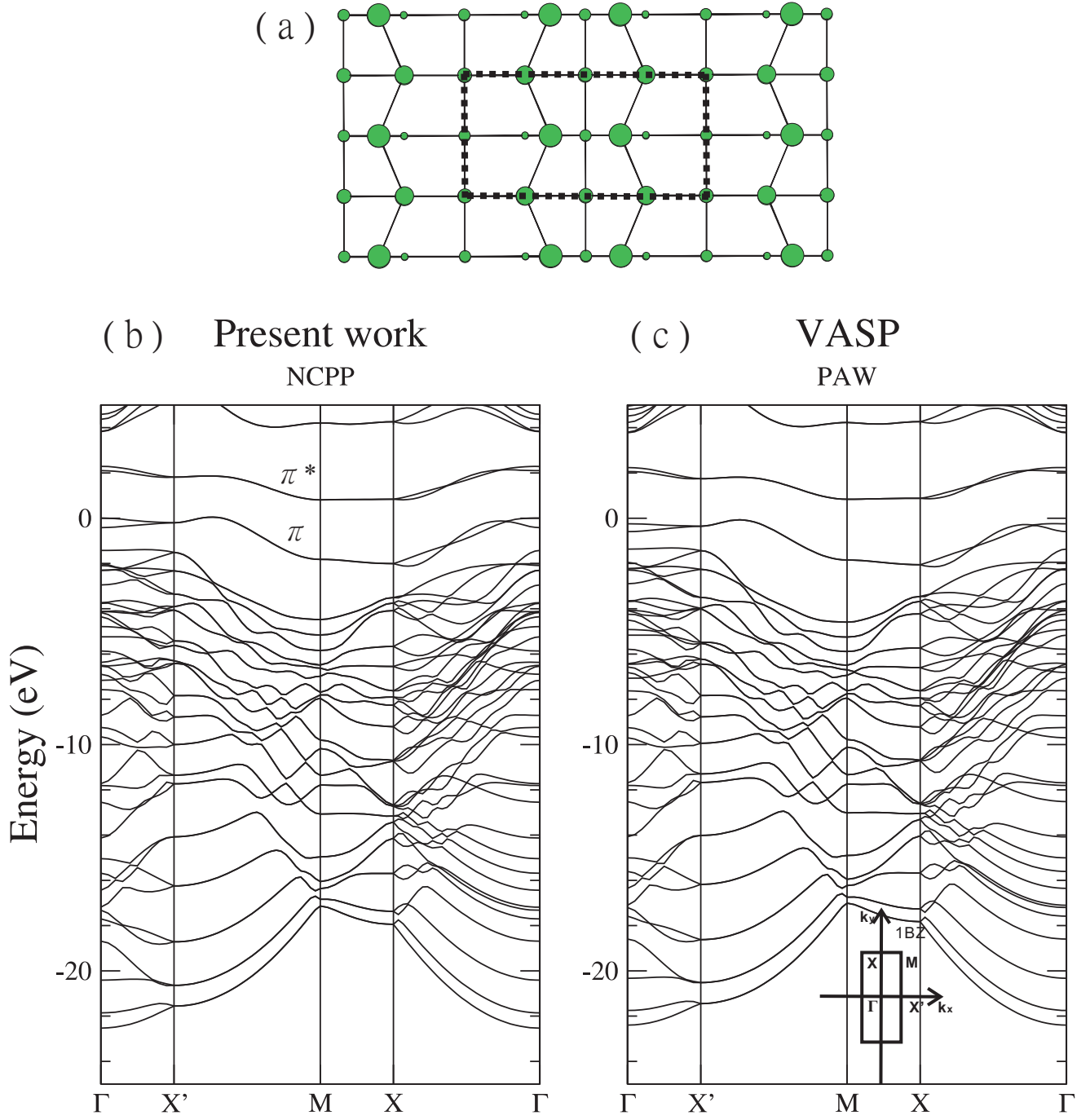


FIG. 4:

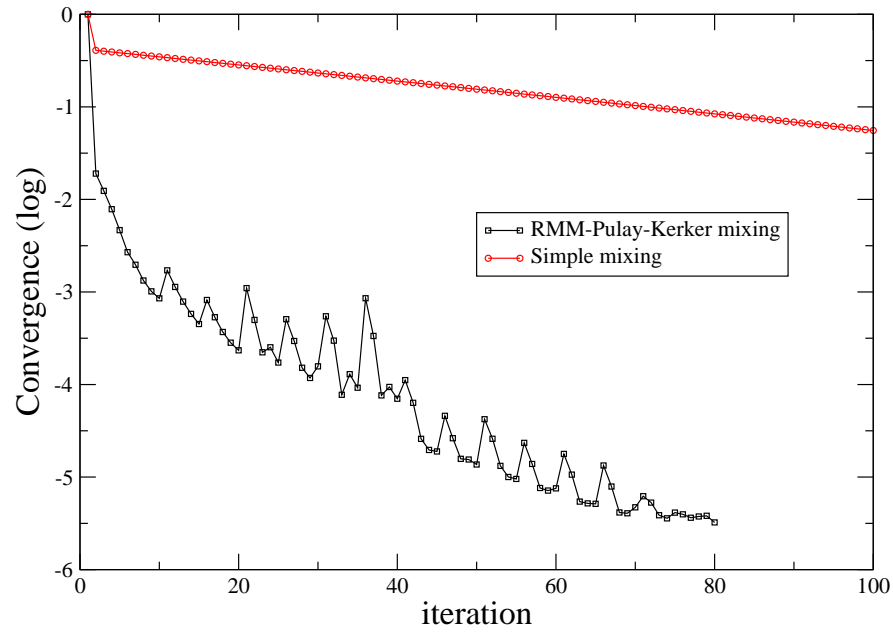


FIG. 5:

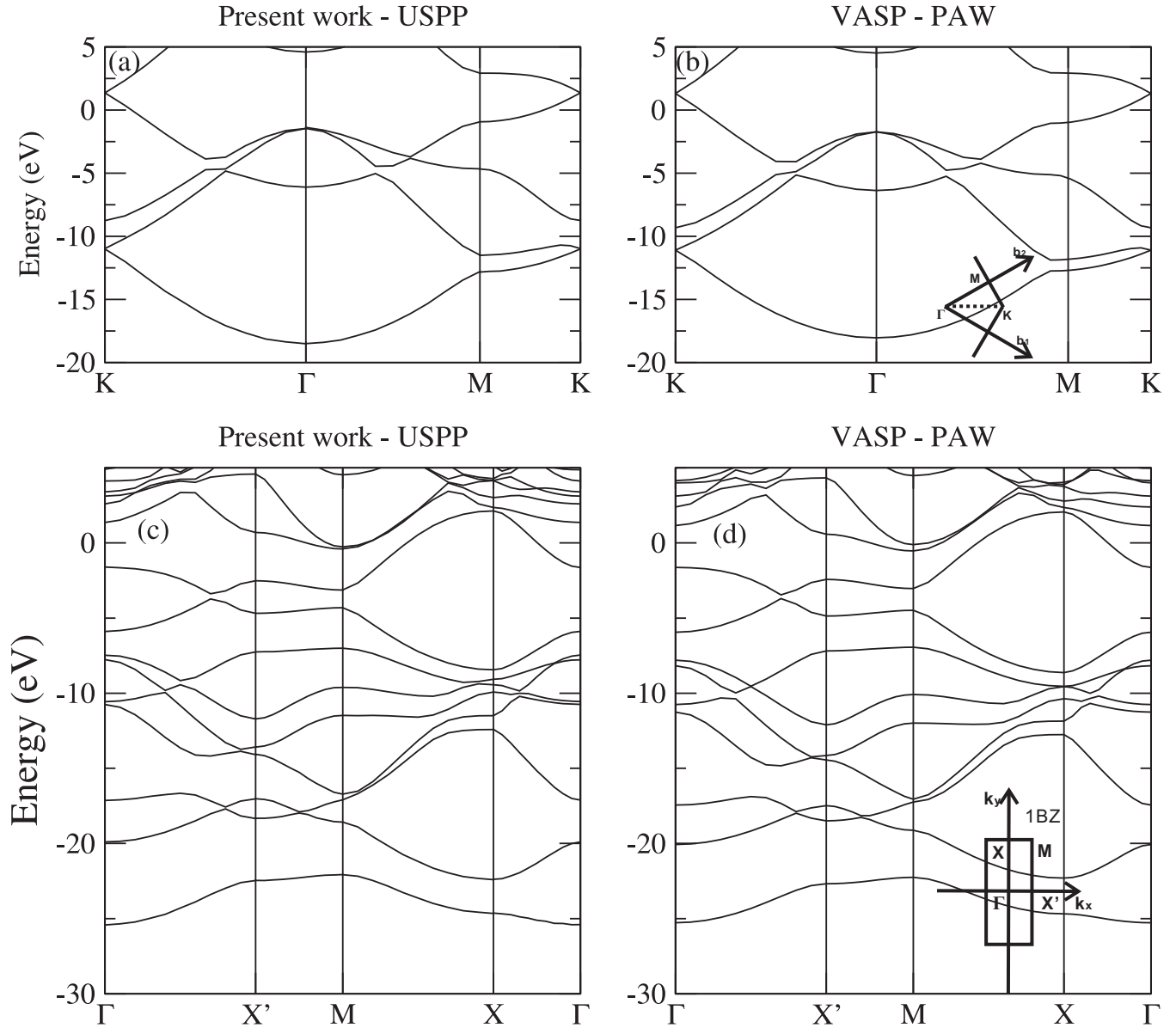


FIG. 6:

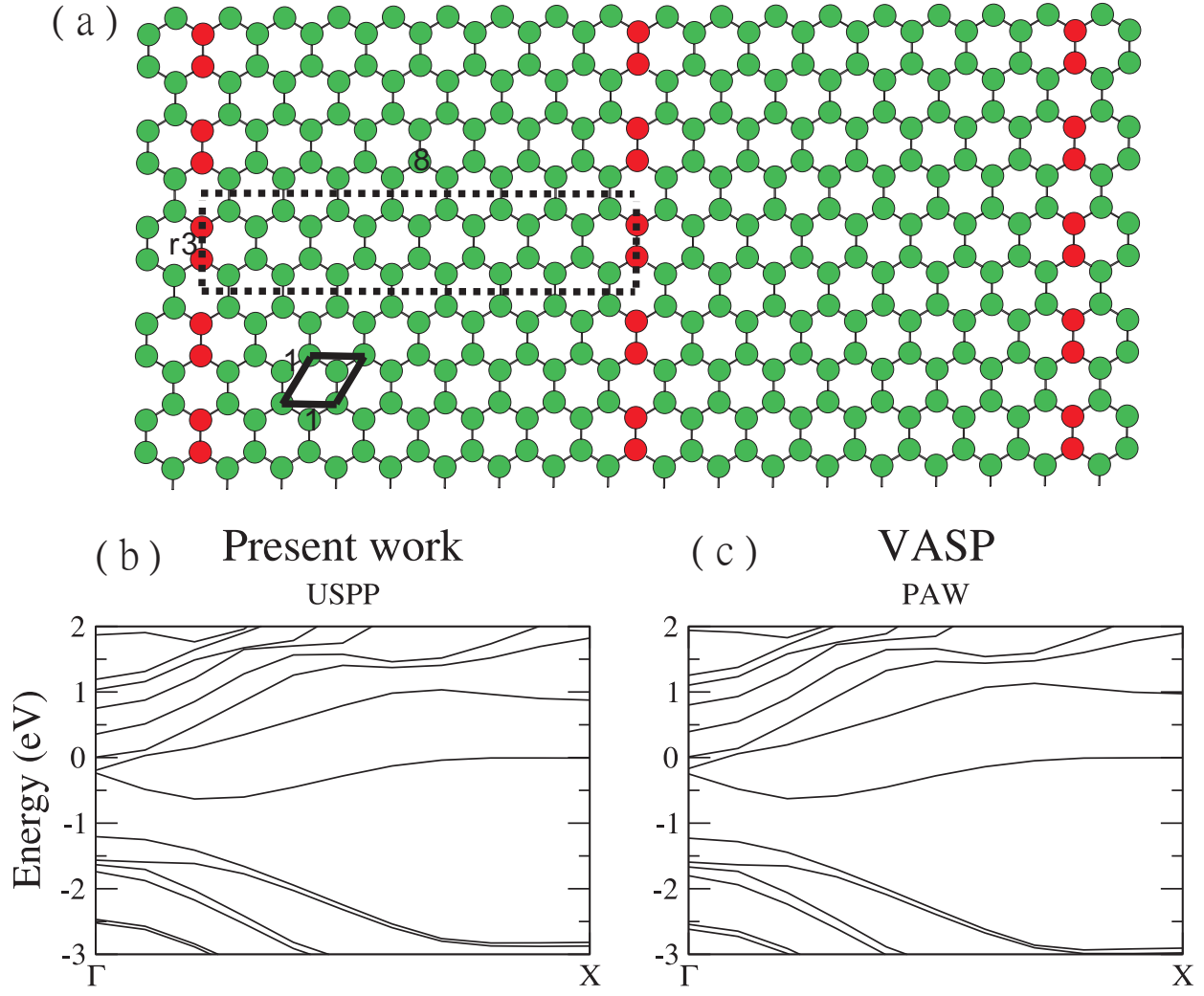


FIG. 7:

